SOME PROBLEMS OF DIOPHANTINE APPROXIMATION: THE SERIES $\Sigma e(\lambda_n)$ AND THE DISTRIBUTION OF THE POINTS $(\lambda_n \alpha)$

By G. H. Hardy and J. E. Littlewood

TRINITY COLLEGE, CAMBRIDGE, ENGLAND

Communicated by E. H. Moore, December 5, 1916

1. In our previous writings on the subject of Diophantine approximation, which we refer to in a short note published in the October number of these PROCEEDINGS,[1] we alluded in several places to a series of further results which, we hoped, were to form the material for a third memoir in the *Acta Mathematica.* The prosecution of this work was delayed, in the first instance, by our occupation on a long memoir on the theory of the Riemann Zeta-function, now in type and shortly to appear there, and subsequently by other causes; and there is, under present conditions, little hope of its completion in the immediate future. The subject has since been reopened by the appearance of work by other writers,[2] and in particular of a very beautiful memoir by Weyl in the latest number of the *Mathematische Annalen.*[3] This paper contains allusions to our unpublished work: and it seems desirable that we should make some more definite statement than has appeared hitherto of our results and the relations in which they stand to Weyl's.

The main problems which we considered were three.

2. (*a*) The first problem was that of proving that, *if* $e(x) = e^{2\pi i x}$ *and*

$$\lambda_n = \alpha n^k + \alpha_1 n^{k-1} + \ldots + \alpha_k$$

is a polynomial in n with at least one irrational coefficient, then

$$s_n = \sum_1^n e(\lambda_h) = o(n).$$

We may plainly suppose that every α has been reduced to its residue to modulus unity: and there is no substantial loss of generality in supposing the *first* coefficient irrational.

This theorem we enunciated first, in the special case in which $\lambda_n = \alpha n^k$, in our communication to the Cambridge Congress, characterising the proof as 'intricate.' In our second memoir in the *Acta* we discussed in detail the case $k = 2$, using a transcendental method which leads to a whole series of more precise results; and we promised a proof of the more general theorem in the third memoir of the series. Weyl's memoir contains a complete statement and proof, both quite independent of ours, of the theorem in its most general form.

The limitation on the form of λ_n, which appears in the theorem as we stated it, was introduced merely for the sake of compactness of expression and does not correspond to any real simplification of the problem. Our argument indeed depends upon an induction which compels us to consider the problem generally. The most comprehensive result which appears in our analysis is as follows: *given any positive numbers* ϵ *and* η *we can determine* $\nu(\epsilon)$, $N(\epsilon, \eta)$, *and a system of intervals* j, *including all rationals whose denominators are less than* ν, *and of total length less than* η, *so that* $|s_n| < \epsilon n$ *for* $n > N$, *all values of* α *exterior to the intervals* j, *and all values of* $\alpha_1, \alpha_2, \ldots, \alpha^k$. From this result it follows at once that $s_n = o(n)$ for any particular irrational α, and uniformly in $\alpha_1, \alpha_2, \ldots, \alpha^k$.

Weyl's proof and ours are widely different, and each, we hope, may prove to have an interest of its own. The same is true of the deduction of the formula $\zeta(l + it) = o(\log t)$, made by Weyl as well as by ourselves.

3. (*b*) The second principal problem was, to use Weyl's phraseology, that of the 'uniform distribution' (*Gleichverteilung*) of the points (λ_n) where (x) is the residue of x to modulus unity. Suppose that n_l is the number of the first n such points which fall within an interval j of length δ. Then the points are said to be *uniformly distributed* if $n_j \sim \delta n$ for every such interval j. It is plain that a corresponding definition may be given of uniform distribution of an enumerable sequence of points in space of any number of dimensions.

That the points (λ_n) are uniformly distributed when $k = l$ and α is irrational was proved independently by Bohl, Sierpinski, and Weyl in 1909–10. The general result (with the same unessential limitation as to the form of λ_n) was stated by us in our first paper in the *Acta*. Our proof, which has never been published, proceeded on the same lines as that of the thoerem of §2. But Weyl has now established a 'principle' which renders such a proof entirely unnecessary, and which has led him to results in this direction far more comprehensive than any of ours. This 'principle' is expressed by the theorem: *if*

$$\sum_1^n e(m\lambda_h) = o(n)$$

for every positive integral value of m, then the points (λ_n) *are uniformly distributed in* $(0, 1)$. The proof depends on a simple but ingenious use of the theory of approximation to arbitrary functions by finite trigonometrical polynomials; and there is a straightforward generalisation to space of any number of dimensions.

Weyl's 'principle' enables him to deduce, with singular ease and elegance, theorems of 'uniform distribution' from theorems of the character of that of §2, and to generalise them immediately to multidimensional space. It enables him to prove, for example, that *the points whose coordinates are*

$$(n^p \alpha_q) \ (p = 1, 2, \ldots, k;\ q = 1, 2, \ldots, l;\ n = 1, 2, 3, \ldots)$$

where $\alpha_1, \alpha_2, \ldots, \alpha_l$ *is any set of linearly independent irrationals, are uniformly distributed in the 'unit cube' of kl dimensions.* All that we had been able to prove was that the points were *everywhere dense* in the cube.

4. (*c*) Corresponding questions arise in connection with an arbitrary increasing sequence $\lambda_1, \lambda_2, \lambda_3, \ldots\ldots$ Are the points $(\lambda_n \alpha)$, for example, uniformly distributed? The answers to such questions in general involve an unspecified exceptional set of values of α of measure zero, instead of (as when $\lambda_n = n^k$) a specified set such as the rationals; they are, in other words, only 'almost always' true.

In our first paper in the *Acta* we proved quite generally that the set $(\lambda_n \alpha)$ is almost always everywhere dense. The corresponding theorem of 'uniform distribution' we discussed only in one especially interesting particular case, that in which $\lambda_n = a^n$, where a is an integer. The theorem is in this case substantially equivalent to results obtained by Borel,[4] from the standpoint of the theory of probabilities, and by Faber,[5] as a corollary of Lebesgue's theorem that a rectifiable curve has a tangent at almost every point. Our analysis however contains the first direct and general discussion of the problem, and leads to results notably more precise than that of mere uniformity of distribution. These results were afterwards made the subject of important generalisations by Fowler,[2] whose investigations covers all cases in which λ_n increases with tolerable regularity and as fast as an exponential of the type e^{n^ρ}. Weyl's 'principle' enables him to reduce this problem to a study of the series $\Sigma e\,(\lambda_n \alpha)$, and leads him to the following theorem, so far the most general of its kind. *If* $c > 0$, $\delta > 0$, *and* λ_n *increases by at least c whenever n increases from h by as much as h* $(\log h)^{-1-\delta}$, *then*

$$s_n = \sum_{1}^{n} e\,(\lambda_h \alpha) = o\,(n), \tag{1}$$

and the points $(\lambda_n \alpha)$ *are uniformly distributed, for almost all values of* α.

In our second paper in the *Acta* we stated that the equation could, in very many cases, be replaced by the much more precise equation

$$s_n = O\,(n^{\frac{1}{2}+\epsilon}) \tag{2}$$

for every positive ϵ. The publication of Weyl's work had led us to a re-examination of this question and to the following theorems.

A. *If* (i) $\lambda_{n+[n^\epsilon]} - \lambda_n \to \infty$,

(ii) $|a_1|^2 + |a_2|^2 + \ldots + |a_n|^2 = O(n^{1+\epsilon})$,

for every positive ϵ, *then*

$$\text{(iii)} \qquad s_n(\alpha) = \sum_1^n a_h e(\lambda_h \alpha) = O(n^{\frac{1}{2}+\epsilon})$$

for almost all α*'s and every positive* ϵ.

B. *If* (i) *is replaced by* (i′) $\lambda_{n+[n^{\beta+\epsilon}]} - \lambda_n \to \infty$, *where* $0 < \beta < 1$, *then* (iii) *may be replaced by*

$$\text{(iii}'\text{)} \qquad s_n(\alpha) = O\left(n^{\frac{1+\beta}{2}+\epsilon}\right).$$

To these two theorems Weyl's forms a completing third. It should be observed that (i) is certainly satisfied if $\lambda_{n+1} - \lambda_n \geq c > 0$, and in particular if λ_n is always integral, and (ii) if $a_n = O(n^\epsilon)$, and in particular if $a_n = 1$.

If λ_n is an integer, and we separate the real and imaginary parts in the equation (iii), we obtain a theorem concerning a particular system of normal orthogonal functions for the interval (0, 1), viz., the functions $\sqrt{2} \cos 2\pi\lambda_n x$, $\sqrt{2} \sin 2\pi\lambda_n x$. Our argument is then directly extensible to a general orthogonal system, and we are led to a new and interesting proof of Hobson's[6] theorem that *if* $\phi_n(x)$ *is any normal orthogonal* system, and $\Sigma n^\delta |c_n|^2$ *is convergent for some positive* δ, *then* $\Sigma c_n\phi_n(x)$ *is convergent almost everywhere.*

Weyl's hypothesis concerning λ_n asserts, roughly, that the increase of λ_n is appreciably more rapid than that of $(\log n)^2$. It is easy to see that this hypothesis cannot be capable of *much* wider generalisation. For, when $\lambda_n = \log n$, s_n is definitely of order n. It seems probable, too, that the index $\frac{1}{2}(1+\beta)$ of Theorem *B* is the correct one.

5. We conclude by correcting an error in our recent note. The results concerning the special case $\rho = 0$ are stated wrongly. It is not true that, when $\rho = 0$, $f(z)$ and s_n are bounded; all that we can assert is that they are of the forms $O\left(\log\frac{1}{1-r}\right)$ and $O(\log n)$ respectively. That $f(z)$ should be bounded would contradict a general theorem of Fatou,[7] in virtue of which a bounded function must tend to a limit, for almost all values of θ, when $z = re^{i\theta}$ tends to the circle of convergence along a radius vector. The error has no bearing on the general case.

[1] Hardy, G. H., and Littlewood, J. E., these PROCEEDINGS, 2, 1916, (583–586).

[2] Berwick, W. E. H., *Mess. Math., Cambridge*, **45**, 1916, (154–160); Fowler, R. H., *London, Proc. Math. Soc.*, (Ser. 2), **14**, 1915, (189–207); Kakeya, S., *Tôhoku Sci. Rep. Imp. Univ.*, **2**, 1913, (33–54) and *Ibid.*, **4**, 1915, (105–109).

[3] Weyl, H., *Math. Ann., Leipzig*, **77**, 1916, (313–352); see also *Göttingen Nachr. Ges. Wiss.*, 1914, (234–244).

[4] Borel, E., *Palermo, Rend. Circ. Mat.*, **27**, 1909, (247–271); see also notes to Borel, E., *Leçons sur la théorie des fonctions*, 2d. ed., Paris.

[5] Faber, G., *Math. Ann., Leipzig*, **69**, 1910, (372–443), especially p. 400.

[6] Hobson, E. W., *London, Proc. Math. Soc.*, (Ser. 2), **12**, 1912, (297–308).

[7] Fatou, P., *Acta Math., Stockholm*, **30**, 1906, (335–400), especially p. 349.

ON MOSELEY'S LAW FOR X-RAY SPECTRA

By Horace Scudder Uhler

SLOANE PHYSICAL LABORATORY, YALE UNIVERSITY

Communicated by B. B. Boltwood, December 23, 1916

While engaged in making interpolations, by the method of least squares, of unknown from known wave-lengths of high frequency spectra I noticed certain systematic deviations from Moseley's law which led me to investigate three interesting questions that have not been previously discussed, probably because the older data did not seem to be sufficiently accurate to justify close mathematical analysis. These questions are: (i) How accurately does Moseley's law reproduce the observed wave-lengths? (ii) What empirical formula will represent the numerical data within the limits of experimental error? and (iii) What is the order of magnitude of the high frequency radiations of elements having small atomic numbers and of which the spectra have not yet been obtained? In the following paragraphs definite answers will be given to questions (i) and (ii), while a tentative solution of the third problem is necessitated by the fact that it involves extrapolation. The wave-lengths used in the computations were taken from the recent papers by M. Siegbahn, W. Stenström, and E. Friman. These data were chosen because they are the latest, they were all obtained in the same laboratory with the same spectrometers, and they constitute the most extensive, accurate and consistent set available.

Moseley's law is that, for any one series (α, β, γ, etc.), the square-root of the frequency of the lines is a linear function of the atomic numbers of the radiating elements. In symbols $\sqrt{\nu_N} = a + b\,N$, where $\nu_N \equiv$ frequency, $N \equiv$ atomic number, a and b are constants for one series. When a and b are calculated by the method of least squares, from the 48 known wave-lengths of the L-α_1 series, extending from zinc ($N = 30$, $\lambda = 12.346$ A) to uranium ($N = 92$, $\lambda = 0.911$ A), the values